Cultural Literacy and Health

John Kelleher

Sometimes when a science notices a conceptual inadequacy in its discourse, the problem is not the rigor of its categories but the categories themselves. When epidemiologists as much as say that they are embarrassed by the limitations of a long-employed concept such as race,[1,2] this may be a hint that something more is needed than a better formalization of familiar categories.

Specifically, neither "race" nor "socioeconomic status" nor "education" may adequately explain intellectual and communicative access, or how it can affect health. "Cultural literacy" may be a concept that can help to fill that breach. Cultural literacy is a term that roughly means "a good grasp of general mainstream knowledge." (The specific term "cultural literacy" is not necessarily familiar to all researchers who nonetheless employ the concept.) Epidemiologists are invited to consider the possibility that cultural literacy—a concept derived not from the social sciences but from the cognitive sciences—may be a practical and productive tool for studying the links between education (or race, or social class) and health.

Cultural literacy is a theory proposed to answer the question: what causes "literacy"? The theory of cultural literacy rests on the "knowledge-competence principle," a fundamental concept in cognitive research. In essence, the principle emphasizes the importance to intellectual competence of relevant prior knowledge. General discussions of the knowledge-competence principle can be found in textbooks on cognitive psychology.[3]

Comprehension is not an abstract "skill." If our brains lack relevant prior knowledge, they magically no longer demonstrate the "skill" of "comprehension." According to the theory of cultural literacy, both oral and written communications entail *mutual* comprehension between parties. Consequently, "literacy" must depend on *mutual* relevant prior knowledge. Thus, to be "culturally literate" means to *share* much of the same broad, mid-level knowledge that other literate members of a society have acquired, either in school or from their reading. The theory of cultural literacy asserts that a solid grasp of broad knowledge about the general mainstream culture *causes*—and to a large extent *is*—increased communicative access to the culture itself.

". . .if cultural literacy is real, then whatever anyone thinks of it, it will sit there calmly, having effects, patiently waiting for epidemiologists to account for them. . ."

The theory was first formulated in the mid-1980s by Hirsch and was widely disseminated via a book-length treatment titled *Cultural Literacy*.[4] The word "cultural" in this context arises because the theory proposes that communicative access (literacy) requires knowledge that is broadly shared. By the mid-1990s, a number of researchers had provided evidence in support of Hirsch's general argument.[5–7,8(p12)]

The words "causes" and "is" are strong words in epidemiologic circles. Epidemiologists should consider well the implications of their use in the description of the theory provided above. Now 15 years since first publication of a book-length treatment of the theory (an eternity in any modern experimental science), not a single competent investigator has cited or performed cognitive research that even quibbles with the general argument of the theory of cultural literacy.

In fact, a test of cultural literacy was shown to provide by itself all the information the U.S. Congress requested from the National Adult Literacy Survey.[9] Hofstetter *et al.*[10] have demonstrated that in a study population demographically similar to the United States, cultural literacy not only exists, but also matters quite a lot.

The concept of cultural literacy is inherently statistical, rather than normative or prescriptive, but it will

Editors' note: *An invited response on this commentary appears on page 500.*

Address correspondence to: John Kelleher, 70 B Cushing Street (rear), Cambridge, MA 02138; johnkelleheriii@yahoo.com

DOI: 10.1097/01.EDE.0000024136.78650.04

acquire a relatively precise and stable meaning in any given complex society. Our brains bluntly refuse to communicate without *shared* frames of reference. "Mainstream knowledge" is not merely a social construct—thus dispensable, or inherently sinister—but rather an indispensable *cognitive* artifact of our brains at work in complex, diverse modern societies. Our brains insist on it. People are diverse, specialties are diverse; nonetheless, brains require *shared* frames of reference. Utopia itself will have "mainstream knowledge."

In Utopia, of course, any two people may know many different things, because of their different specialties and home backgrounds. But both will share the mid-level mainstream frames of reference—cultural literacy—that make effective communication statistically more likely.

Epidemiologists will find that the theory of cultural literacy is highly controversial in the sense that it is probably inconsistent (even dramatically inconsistent in some cases) with widespread academic, philosophical, and "folk" (homemade) theories of knowledge and communication. It is also inconsistent with political, social, and educational premises that implicitly depend on such theories. On the other hand, if cultural literacy is real, then whatever anyone thinks of it, it will sit there calmly, having effects, patiently waiting for epidemiologists to account for them.

Cultural literacy is part of the basic communications structure of a diverse, modern, knowledge-based society. Thus, questions related to infrastructure are relevant to a discussion of cultural literacy. As an analogy, owning a telephone is probably a good thing for you as an individual, but even if you own a telephone, it still matters whether you're able to dial 1%, 50%, or 100% of your fellow citizens. It matters whether telephones are abundant in some places and scarce in others, or whether variations in the quality of telephones make some communications easy and others arduous or practically impossible. In the same way, having an extensive grasp of mid-level mainstream knowledge may be a good thing for you individually, but at the same time the overall level and distribution of cultural literacy in your society will still matter, independent of your own personal knowledge. Epidemiologists may have to consider such "infrastructure" effects on the public health and on health care costs.

Of course, people's individual levels of cultural literacy may affect their own ability to understand the requirements of good health, their access to health care, and their ability to pay for care. For example, more knowledgeable parents in industrializing Europe made greater use of new medical advances. This in turn was associated with lower morbidity and mortality rates in their children.[11]

Another example: low-literate Americans have poor ability to understand and properly use healthcare resources, which may both decrease their use of helpful resources and increase their unnecessary use of those resources.[12] Client-centered remedies include distributing greatly simplified health messages.

By implication, cultural literacy, which causes high general literacy, is protective in this situation. This is so even though, at an individual level, those with broad general knowledge will not automatically know *all* information relevant to their use of and access to healthcare.

Another implication of the theory of cultural literacy is that a higher level of cultural literacy increases further intellectual access—on average, it makes you a quicker study. The broader your existing knowledge, the more likely it is that you already know something that will make it easier to learn something else.

In the United States, at least, broad general knowledge appears to be strongly associated with educational attainment (about twice as strongly as socioeconomic status is associated with educational attainment).[13] It is also associated with power—"access" in the most basic sense. Moreover, cultural literacy can be measured by simple checklists that can be used over the telephone.[9,10] Using random-digit dialing in San Diego, California, a study found that possession of a broad range of declarative knowledge about the mainstream culture was associated with achieving and manifesting power (as measured by occupation, income, and political activity).[10] This held even after controlling for age, education, and ethnicity.

Epidemiologists should probably not think about cultural literacy as if it were a category (like socioeconomic status), but more as a continuous variable that can be assessed for each individual. It is something specific enough to measure, and it is well connected to empirically validated and conceptually rigorous theory (albeit a theory developed outside of epidemiology).

Given the above, a clear epidemiologic question can now be posed in a manner amenable to empirical investigation: is there an association between cultural literacy and good health? If so, what might be the causal pathway? Does cultural literacy produce higher income and increased power, with the health benefits that follow? Or does cultural literacy allow persons better access to health care through better communication with health professionals, better understanding of medical directions, more effective advocacy for their own needs, or increased motivation to modify behaviors?

Epidemiologists should also note the possibility that increased cultural literacy may not produce better health. To assume that the consequences of cultural literacy will be purely positive requires at least three other assumptions, all highly debatable. One, it assumes that people will never use their powers to obtain things that worsen their health. Two, it assumes experts are

always right, and therefore increased access to experts leads to better health. Three, it assumes that all the information comprising cultural literacy is actually conducive to better health. Finally, there may be circumstances in which cultural literacy is actually punished, to the detriment of those possessing it (see below).

A brief excursus into "years of education" may introduce epidemiologists to a bit of the subtlety behind ideas only broadly sketched here. "Years of education" is probably not an adequate marker for cultural literacy (maybe especially in the United States). For example, when Ferguson controlled for actual educational attainment (as measured on the Armed Forces Qualification Test) rather than years of education, the income gap between blacks and whites in the United States decreased from 16 per cent to less than 5 per cent.[14] Which is to say, the current American way of schooling at the primary and secondary levels may be differentially unhelpful to the educational achievement of African-Americans.

A prominent initial source of most individuals' cultural literacy is school. Schooling is a specific measurable entity with empirical and theoretical links not only to communicative access but also to general work-related competence.[15] This can account for variance in income among American blacks and whites at least as well as socioeconomic status or race—categories whose conceptual clarity, proper modern scientific foundation, theoretic richness, explanatory value, and formal rigor epidemiologists have recently begun to question.

This is not to argue that other categories should be eliminated. Studies of cultural literacy often also employ the concepts of income and race. The question is simply which categories are most predictive of the outcome under study. In Ferguson's study of income,[14] the category of "race" was not the most pertinent in explaining the observed variance.[16]

Should studies of health measure cultural literacy? Could cultural literacy more precisely account for some health effects previously attributed to other factors? Because of differentially poor schooling, could it be that race and socioeconomic status sometimes actually serve as crude markers for lack of general mainstream knowledge? At the moment, questions such as these are almost entirely open.

In some modern societies, there may well be ethnic and class differences in cultural literacy that are relevant to health. By itself, lower cultural literacy does not indicate an absolute "knowledge deficit," but only a relative deficiency in mainstream knowledge. Moreover, mainstream knowledge, although crucial for intellectual and communicative access, is neither automatically "better" nor even necessarily correct. What to do? Probably, there is no free lunch. At the individual level, real people still are well or ill, and need (and sometimes mistakenly misuse) health care resources, provided by distinctly non-utopian societies.

At the system level, Hirsch[4,8] has argued that high mean levels and narrow distributions of cultural literacy do not just happen; heterogeneous societies desiring these must painstakingly aim for both, in both theory and practice. Certain national models of universal schooling dramatically reduce ethnic (*eg*, Third World immigrant) and class variance in general academic knowledge while also producing high mean levels.[17,18] Adult education and literacy programs can succeed,[19] and mothers attending them are better at nurturing literacy in their children.[20] Plain old lifelong general reading increases cultural literacy.[7]

For completeness, it is worthwhile to make a distinction between the association of cultural literacy with communicative access, and the association of communicative access with power and health. Although the theory of cultural literacy proposes that firm possession of the shared mid-level background knowledge of the mainstream culture *per se* causes communicative access to that culture, population increases in cultural literacy *per se* do not create jobs, democratic participation, and health. Hence, in a given society, increased cultural literacy could merely create large masses of disaffected people whose only increased "power" would be the ability to offer more sophisticated accounts of why they are without jobs, have no say, and suffer ill health. To take an even more extreme example, in the 1970s, the Communist Khmer Rouge "tried to turn Cambodia into a self-sufficient, agrarian utopia. They forced people to move from cities to the countryside and murdered the educated and the skilled."[21] In Cambodia under the Khmer Rouge, to be culturally literate was to be murdered.

Given the examples here, the firmly causal relation between increased mainstream knowledge and increased mainstream literacy does not of necessity produce income, power, and health; such an association can exist and even be causal, but only within the possibilities created by larger societal beliefs and structures.

The theory of cultural literacy may give epidemiologists a more precise and scientifically rigorous way to think theoretically about the issue of intellectual and communicative access. Recent research may provide some hints about how to undertake its practical empirical investigation.[9,10] Interested epidemiologists will find some overlap of knowledge and interest with colleagues in other fields (for example, political scientist Hofstetter[10] is also adjunct professor in a school of public health). Good luck to all the brave pioneers of this new conversation.

Acknowledgments

I thank Mary Barger, Eugene Declercq, David C. Geary, E. D. Hirsch, C. Richard Hofstetter, and Thomas G. Sticht, who read drafts of the paper and contributed valuable comments.

References

1. Kaufman JS, Cooper RS, McGee DL. Socioeconomic status and health in blacks and whites: the problem of residual confounding and the resiliency of race. *Epidemiology* 1997;8:621–628.
2. Morgenstern H. Defining and explaining race effects (editorial). *Epidemiology* 1997;8:609–611.
3. Baddeley A. *Human Memory: Theory and Practice*. Needham: Allyn and Bacon, 1998;125–143.
4. Hirsch ED. *Cultural Literacy: What Every American Needs to Know*. Boston: Houghton Mifflin, 1987.
5. Kosmoski GJ, Gay G, Vockell EL. Cultural literacy and academic achievement. *J Exp Educ* 1990;58:265–272.
6. Pentony JF. Cultural literacy. *Adult Basic Educ* 1997;1:39–45.
7. Stanovich KE, West RF, Harrison MR. Knowledge growth and maintenance across the life span: the role of print exposure. *Dev Psychol* 1995;31:811–826.
8. Hirsch ED. *The Schools We Need and Why We Don't Have Them*. New York: Doubleday, 1996.
9. Sticht TG, Hofstetter CR, Hofstetter TG. Assessing adult literacy by telephone. *J Lit Res* 1996;28:525–559.
10. Hofstetter CR, Sticht TG, Hofstetter TG, Huie C. Knowledge, literacy, and power. *Communic Res* 1999;26:58–80.
11. Geary DC. Evolution and proximate expression of human paternal investment. *Psychol Bull* 2000;126:55–77.
12. Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs, American Medical Association. Health literacy: report of the Council on Scientific Affairs. *JAMA* 1999;281:552–557.
13. Lubinski D, Humphreys LG. Incorporating general intelligence into epidemiology and the social sciences. *Intelligence* 1997;24:159–201.
14. Ferguson RF. Shifting challenges: fifty years of economic change towards black-white earnings equality. *Daedalus* 1995;124:1, 37–76.
15. Scribner BLS, Smith DA, Baldwin RH, Phillips RL. Are smart tankers better? AFQT and military productivity. *Armed Forces Soc* 1986;12:193–206.
16. Jencks C, Phillips M, eds. *The Black-White Test Score Gap*. Washington DC: Brookings Institute, 1998;445, 489–494.
17. Erikson R, Jonsson J, eds. *Can Education Be Equalized? The Swedish Case in Comparative Perspective*. Boulder: Westview Press, 1996.
18. Hirsch ED. Equity effects of very early schooling in France. Available at: http://www.coreknowledge.org/CKproto2/Preschool/FrenchEquity.htm. Accessed 22 June 2002.
19. Sticht TG. *The Adult Education and Literacy System (AELS) in the United States: Moving from the Margins to the Mainstream of Education*. Washington DC: U.S. Department of Education, Division of Adult Education and Literacy, 2000.
20. Van Fossen S, Sticht TG. *Teach the Mother and Reach the Child: Results of the Intergenerational Literacy Action Research Project of Wider Opportunities for Women*. Washington DC: Wider Opportunities for Women, 1991.
21. *PBS NewsHour, 18 June 1997* (transcript). Alexandria, VA: Public Broadcasting Service, 1997.

Whad'Ya Know? Another View on Cultural Literacy

Jay S. Kaufman

"Whoever undertakes to set himself up as a judge of Truth and Knowledge is shipwrecked by the laughter of the gods."
– Albert Einstein (1879–1955)

The association between education and good health has been observed widely, and a large number of epidemiologic studies over the past decades have included attained educational level as an exposure or covariate.[1] The basis for this association is no doubt multifaceted, including connections (education provides access to elite social networks), credentialing (degrees confer entrée to positions of power and authority), confounding (those with the wherewithal to obtain greater quantity and quality of education also have the wherewithal to obtain a greater quantity and quality of other goods), and content (education actually confers some useful information that helps people gain advantages in life).[2] The essay by John Kelleher[3] in the current issue of this journal suggests an improved exposure assessment method for the latter of these components of education, the useful content of a quality education itself. This suggestion relies heavily on the notion of "cultural literacy" introduced in the 1980s by educator E. D. Hirsch. Cultural literacy is the theory that there are certain things that everyone in a modern society ought to know, and that it is the possession of these various pieces of knowledge that confers to individuals the means to understand, communicate, and succeed—both socially and materially.[4–6]

From the Department of Epidemiology, University of North Carolina School of Public Health and Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC.

Address correspondence to: Jay S. Kaufman, Department of Epidemiology (CB #7435), University of North Carolina School of Public Health, McGavran-Greenberg Hall, Pittsboro Road, Chapel Hill, NC 27599-7435; Jay_Kaufman@unc.edu

DOI: 10.1097/01.EDE.0000024137.49946.F5

Hirsch and colleagues not only suggested that there were words and ideas that were central to our common American culture, but they also drafted lists of these items and they devised tests, based on these lists, in order to assess an individual's attained level of cultural literacy.[7] Hirsch's 1987 book contained the first such list, including 63 pages devoted to nearly 5000 names, dates, aphorisms, and concepts that Hirsch and his collaborators asserted everyone ought to know. The methodology for generating this list has remained somewhat mysterious, however. Hirsch explained merely that "more than one hundred consultants reported agreement on over 90 percent of the items listed,"[4(p146)] without ever explaining how such agreement was ascertained (*eg,* did each consultant generate his or her own list independently?), nor indeed how the consultants themselves were chosen.[8(pp80–81)] Furthermore, Hirsch alluded to rejecting some items from the list because they were in fact too commonly familiar, and thus failed to discriminate effectively between the truly literate and the masses.[4(p146)] He also made the list occasionally proscriptive rather than descriptive by including items, especially from the natural sciences, that were generally unfamiliar to even the more culturally literate, but which he felt *should be* familiar.[4(p148)]

"Hirsch's list is the disease for which it claims to be the cure," argued Neil Postman. "[T]hat is to say, its arbitrariness only demonstrates the futility of trying to do what he wants to do."[9(p121)] Indeed, the original list and its variants have been savaged by some for their inexplicable capriciousness and by others for their overt bias. For example, in the arena of entertainment, Hirsch declared that all Americans ought be familiar with P. T. Barnum, Greta Garbo, and the Barrymores, while omitting Orson Welles, Mickey Mouse, and Miles Davis. Likewise, seven leading businessmen and industrialists were listed, but no counterpart labor leaders such as Samuel Gompers or Jimmy Hoffa. Essential music in our common culture includes, according to Hirsch, "Yankee Doodle" and "White Christmas," but not "Louie Louie" or "Satisfaction." This led Dan Fleming to conclude that "[b]y his list, Hirsch clearly reveals a disdain of the modern world."[10(p109)] In addition to avoiding the contemporary, Fleming noted that Hirsch also avoided the nonwhite, for example by omitting from the list Mexico City (the largest city in the world), while including Stuttgart and Hamburg, or by omitting Mexico and Kenya but including Luxembourg.

Similar observations were made by others, including Herbert Kohl, who concluded that Hirsch thought a culturally literate individual was "apparently a university-educated European-American, most likely male, who speaks in platitudes and has a passing acquaintance with words drawn from the sciences, the humanities and the arts."[11(p457)] To illustrate his point, Kohl excerpted items beginning with the letter "P" with which Hirsch asserted everyone ought to be familiar:

"*perfectibility of man, periodic table of the elements, pax romana, pay the piper, pearl of great price, peeping Tom, Peloponnesian War, penis envy, penny saved is a penny earned, persona non grata, Peter the Great, Phi Beta Kappa, philosopher king, photoelectric effect, plate tectonics, Pickwickian, Planck's constant, play second fiddle, pogrom, proof of the pudding is in the eating, and Pyrrhic victory.*"

Though I have a doctoral degree, I confess to a little bit of confusion about "Pickwickian." However embarrassing it is to admit this as a scientist, I am also not sure how well I could explain the photoelectric effect, plate tectonics, or Planck's constant, nor have I ever heard "pearl of great price" used in a conversation, ever. The question remains whether one qualifies as culturally literate merely from having encountered these words, or whether one must actually understand what they mean, and if so, just how deep an understanding is required.[12] Despite my own rudimentary grasp of many of these concepts, I can at least take solace in the firm belief that few contemporary leaders of government or industry, or others of substantial wealth and influence in our society, could do much better with Planck's constant than I could.

"Essential music in our common culture includes, according to Hirsch, 'Yankee Doodle' and 'White Christmas,' but not 'Louie Louie' or 'Satisfaction'."

Kohl[11] also noted that the list tended to downplay words and phrases that relate to subgroup thinking and non-Western cultures, and listed examples of words and phrases that Hirsch opted to omit such as "peace activists, pesticides, political prisoner, potlatch, premenstrual syndrome, prison, prophylactic, prostitution, pueblo and prime time," not to mention "prick, piss, putz, pussy, patronize, palimony, prissy, putsch, pig, profligate, play politics, play the field, and play into one's hands." Who is to judge that these words are not equally at the center of our culture as those chosen by Hirsch? I certainly do better on Kohl's lists than on Hirsch's, despite generous helpings of elite education. This example highlights a certain arrogance in making such lists in the first place, presuming that Hirsch or any other individual is qualified to speak for what is essential for the rest of us to know. A reasonably objective or rational methodology

for constructing such a list has never, to my knowledge, been suggested by any of the proponents of cultural literacy. Perhaps an even more damning indictment of the arrogance involved in drafting such lists is that Hirsch and his colleagues made so many embarrassing factual errors.[13] For example, they asserted that HIV was unknown in 1986, that President Roosevelt died after Nazi Germany surrendered, that Marx Brothers' films featured a sibling named "Gummo", and that Khrushchev sent troops into Poland in 1956 (it was Hungary). Among other gaffes, they also provided incorrect definitions or explanations of "the national debt," IQ scores, and "The Dow-Jones Average."

More importantly, in a diverse society there is simply no monolithic national culture, and no single set of facts or terms that everyone ought to know in order to succeed materially or socially.[8(pp71–72),14] A good example of confusion over this point occurs in Kelleher's essay when he asserts that ". . .in the 1970s the Communist Khmer Rouge . . . forced people to move from cities to the countryside and murdered the educated and the skilled. In Cambodia under the Khmer Rouge, to be culturally literate was to be murdered."[3(p499)] The crucial point here is that, in fact, many kinds of knowledge exist in a society, and only one kind was targeted by the Khmer Rouge. The regime killed or exiled those who had a Western education, who spoke French, or who were cultivated in matters urbane, cosmopolitan, or bourgeois. There are many kinds of knowledge that are necessary to be a successful farmer in rural Cambodia: an intimate knowledge of various plants and domesticated animals, for example. People proficient in these sources of knowledge were not targeted. The author mistakenly implies that Western cultural literacy, which is what the Khmer regime attacked, is equivalent to cultural literacy in general.

What happened in Cambodia in the 1970s was, in fact, a war between subcultures of knowledge within a complex society, with the Khmer Rouge regime using the coercive power of the state to root out and destroy one such subculture. The United States in the 1980s also witnessed a war between subcultures of knowledge, although ours was, thankfully, decidedly less murderous. Hirsch's work played a prominent role in these "culture war" debates, as did the work of Allan Bloom,[15,16] with which Hirsch is often linked. Bloom was at least more open about his cultural biases. One only needs to note, for example, Bloom's revulsion over the way Louis Armstrong's "smiling face" singing Kurt Weill's "Mack the Knife" symbolized the ignorant Americanization of a purer German culture.[15(p151),17] In Hirsch's work, the cultural chauvinism is perhaps more subtle, but no less readily apparent. It is revealing, for example, that in 1987 Hirsch should include on his list Rhodesia but not Zimbabwe, Ceylon but not Sri Lanka, and Peking but not Beijing.[10(p108)]

Even setting aside the thorny theoretical issues, the practical constraints are daunting. Kelleher reassures the reader that the construct can be measured. Although published empirical work has indeed relied upon adaptations of Hirsch's lists,[18,19] it is by no means clear that these simple checklists capture a quantity that we could assume to be cultural literacy in the sense described by Kelleher. Because Hirsch vacillated between proscriptive and descriptive strategies in the selection of items, no gold standard appears to exist for the validation of any derived scale.[8(pp73–75)] Therefore, although checklists have been administered and scores computed, this exercise provides no basis at all for Kelleher's confident assertion that cultural literacy is "something specific enough to measure,"[3(p498)] only that self-reported familiarity with items on Hirsch's lists is something specific enough to measure.

There is no disputing the observation that the epidemiologic measure of years of completed formal education is a hodgepodge variable that involves numerous components, some causal and some spurious, and that this variable is therefore deficient in a theoretical sense. Thus, the idea of capturing knowledge as one specific causal component of this quantity is an attractive one, especially if the content of educational experience is a salient factor in determining health status. The theoretical and practical barriers to doing so appear to me to be insurmountable, however, and the Kelleher commentary,[3] although raising this interesting possibility, has failed to provide any indication that we can actually do better in this regard. Although Kelleher cites empirical studies in which individuals responded to checklists of items deemed to be indicative of cultural literacy, one should remain highly suspicious of these lists until there is a full accounting of the methodology for their construction and some formal validation of their capacity to capture the unified latent quantity that we might reasonably call cultural literacy in any general sense.

References

1. Kitagawa EM, Hauser PM. *Differential Mortality in the United States: A Study in Socioeconomic Epidemiology*. Cambridge: Harvard University Press, 1973.
2. Yen IH, Moss N. Unbundling education: a critical discussion of what education confers and how it lowers risk for disease and death. *Ann N Y Acad Sci* 1999;896:350–351.
3. Kelleher J. Cultural literacy and health. *Epidemiology* 2002;13:497–500.
4. Hirsch ED. *Cultural Literacy: What Every American Needs to Know*. Boston: Houghton Mifflin, 1987.
5. Hirsch ED. *The Dictionary of Cultural Literacy*. Boston: Houghton Mifflin, 1988.
6. Hirsch ED, Rowland WG, Stanford M, eds. *A First Dictionary of Cultural Literacy: What Our Children Need to Know*. Boston: Houghton Mifflin, 1989.
7. Goldberg MF. Doing what works: An interview with ED Hirsch, Jr. *Phi Delta Kappan* 1997;79:83–85.

8. Smith BH. Cult-lit: Hirsch, literacy, and the "national culture." *South Atl Q* 1990;89:69–88.
9. Postman N. Cultural literacy. *The Atlantic* 1989;264:119–124.
10. Fleming DB. Beware of those bearing lists. *N ASSP Bull* 1992;76: 105–109.
11. Kohl H. Rotten to the Core. *The Nation* 1992;254:457–462.
12. Shenker J. The well-stocked brain (book reviews). *NY Times (Print)* 17 December 1989;Section 7:22, column 1.
13. Seligman D. Getting it wrong. *Fortune* 1993;128:169–170.
14. Klinghoffer D. Trivial pursuit: the pros and cons of cultural literacy. *Natl Rev* 1999;51:61.
15. Bloom A. *The Closing of the American Mind: How Higher Education Has Failed Democracy and Impoverished the Souls of Today's Students*. New York: Simon and Schuster, 1987.
16. Urban WJ, reviewer. *J Am Hist* 1988;75:869–874. Review of: Bloom A. *The Closing of the American Mind: How Higher Education Has Failed Democracy and Impoverished the Souls of Today's Students*; Hirsch ED Jr. *Cultural Literacy. What Every American Needs to Know*; Ravitch D, Finn CE Jr. *What Do Our 17-Year-Olds Know? A Report on the First National Assessment of History and Literature*.
17. Seaton J. Cultural conservatism, political radicalism. *J Am Cult* 1989;12:1–9.
18. Pentony JF. Validity and reliability of the Cultural Literacy Test. *Psychol Rep* 1996;78:1027–1033.
19. Bart WM. On the relationship between cultural literacy and scholastic achievement among university students in secondary education. *Psychol Rep* 1998;82:841–843.

Validation Studies: Bias, Efficiency, and Exposure Assessment

Nilanjan Chatterjee and Sholom Wacholder

Measurement error is the bane of epidemiologic studies of diet, behavior, and environmental factors. Even molecular assays—whether for simple genotypes or for complicated biochemistry—are not immune. In this issue, Stürmer *et al.*[1] make some useful observations on methods to reduce or eliminate the bias from measurement error. Herein, we discuss the general problem of measurement error and comment on the current state of epidemiologic methods to mitigate its effect.

Errors in variables (the term used by statisticians) lead to distorted estimates of effect and to underpowered or biased tests. The impact of errors is well understood, at least regarding the direction of the estimate, in a few important special situations. However, measurement error can introduce unpredictable distortions in many realistic settings.[2,3]

Case-control studies with exposure data or biospecimens collected retrospectively are particularly vulnerable to poor measurement of exposure. *Nondifferential* measurement error is often a consequence of exposure information that is collected long after the exposure occurs. *Differential* error can arise when symptoms or treatment of disease affect a biomarker (other than germ-line DNA) or, along with knowledge of the presence of disease, influence response to questionnaires.

Internal validation studies can be used to reduce the impact of measurement error. An error-prone exposure measurement Z is collected from everyone in the main study. A more accurate but more expensive measurement X is also available, in principle, for everybody. However, owing to cost or practical considerations, X is collected only on a *validation sample*, consisting of small subsets of cases and controls selected randomly. Clearly, the risk parameter associated with X could be estimated unbiasedly but quite imprecisely with a *complete case estimator* (CCE) that discards all the imperfect but informative Z measurements from subjects not in the validation sample. The *regression calibration* (RCE) and *semiparametric efficient estimators* (SPE) exploit the imperfect measurements Z from individuals who were not included in the validation sample to obtain a more efficient estimate of the risk parameter.

What is the basic principle behind these more sophisticated "bias correction" methods that use both sets of measurements? The validation sample reveals the relation between Z and X. Based on this relation, a probabilistic distribution of X can be inferred from Z for subjects with unknown X. SPE and RCE use different ways of predicting X from Z; they make different tradeoffs between stronger assumptions about the structure of the error and greater reliance on the validation data itself. RCE also requires an additional assumption that the measurement error be small.

From the Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD.

Address correspondence to: Nilanjan Chatterjee, 6120 Executive Blvd, EPS 8038, Bethesda, MD 20892-7244; chattern@mail.nih.gov

DOI: 10.1097/01.EDE.0000022948.80077.AE

Understanding the basis of the nomenclature can help one understand the distinction. The term regression calibration describes how Z is calibrated to X based on the parametric regression model for X given Z; incidentally, it also evokes the calibration of the regression of interest by data from the validation study. In contrast, SPE predicts the distribution of X given Z from the validation study *nonparametrically*, that is, without imposing a parametric relation between X and Z. It is deemed *semiparametric* because it involves one parametric and one nonparametric component: the parametric component is the regression model for Y given X, and the nonparametric component is the distribution of X given Z. The method is called *efficient* because it is defined as the most efficient among the class of all semiparametric estimators that treat the distribution of X given Z nonparametrically. In other words, SPE predicts the distribution of X given Z from the validation study nonparametrically, whereas RCE requires a specific parametric assumption about the conditional distribution of X given Z, such as the conditional mean of X being linear in Z, or what Stürmer *et al.* call *linear measurement error*.[1]

An analogy can be drawn with the use of a parametric *t*-test or a nonparametric Wilcoxon rank-sum test when testing whether the distributions of a variable are the same in two groups. The *t*-test relies on the assumptions that the variables are normally distributed with equal variance in both groups; it is more efficient when the assumptions hold, but can be quite misleading when they are strongly violated. In contrast, the nonparametric procedure makes no distributional assumption; its superior performance when the *t*-test assumptions are violated compensates for its lower efficiency when the assumptions hold. In bias correction, the precision of RCE under the correctly specified error model contrasts with the robustness of SPE against departures from the model. The nonparametric aspect of SPE allows it to capture more nuances of the relation between X and Z than RCE. Assume, for example, that multidimensional X is observed with nondifferential measurement error in each component and that Z is the corresponding error-prone variable. Even if X and the errors $Z–X$ are not normally distributed and each of the components of X and $Z–X$ are correlated with each other,[2,3] semiparametric estimators remain valid. When the normality assumption is correct and the relation between Z and X is linear, however, SPE is less efficient than RCE because RCE capitalizes on the imposed structure. Stürmer's simulation studies (Table 2) clearly show that for small measurement error, the efficiency gain of RCE over SPE can be substantial if the assumed linear measurement error model holds, but the authors do not explore the consequence of the violation of this assumption.[1]

Differential measurement error poses an additional

In contrast, we usually take for granted the ability of these methods to produce an unbiased estimate, at least under idealized conditions. We view validation designs as tools for minimizing the cost of a study with fixed power or, equivalently, for maximizing the precision of the estimate of the main study parameter with fixed cost. We therefore consider it valuable to consider designs with stratified random sampling (with strata defined jointly by case-control status and error-prone exposure measurements), which can often be substantially more efficient than either simple random sampling or standard case-control sampling.[6–8] For example, oversampling cases so that the numbers of cases and control in the validation stage are equal is clearly a more efficient strategy for CCE, and, not surprisingly, for RCE and SPE as well, than the simple random sampling Stürmer *et al.* considered. Sampling based on Z could further improve the efficiency of the design. For example, if Z is at least moderately correlated with X, oversampling extreme values of Z yields greater numbers of extreme Xs and, therefore, will be more informative for estimating a slope. There need not be any bias attributable to stratified sampling if appropriate statistical methods are used in the analysis stage.

The validation design is a special case of "two-phase stratified study designs," which can provide cost savings in many epidemiologic studies. The two phases are the collection of a set of inexpensive covariates Z for all subjects, followed by the collection of more expensive covariates X at phase 2 on a smaller subsample of subjects selected based on values of Z and case-control status. More generally, two-phase designs can be used profitably to collect information on an expensive exposure of interest,[9] confounder,[6] or effect modifier[10] on a sample of subjects, with the sampling faction varying according to the value of variables available for everyone in the study. Even "old-fashioned" matching and contemporary countermatching[11] can be seen as two-phase strategies,[9] because collection of exposure X depends on the matching variables Z. In the class of two-phase designs, the validation study is special in only one rather trivial way: Z is not included in the risk model, because it is assumed that there is no information about risk of disease attributable to Z that is not contained in X.

So where are we now? The paper by Stürmer *et al.* shows the promise of sophisticated statistical methods for error correction.[1] In general, semiparametric estimators are more flexible and robust than RCE in the presence of poorly understood error mechanisms. When computation of a semiparametric efficient estimator is overly complex, slightly less efficient but simpler semiparametric estimators based on pseudolikelihood methods[12] can be attractive. Software is now available in S-PLUS (MathSoft, Inc., Seattle)[13] for various semiparametric methods of the general two-phase data problem (including validation studies for nondifferential measurement error problems) using the logistic regression model. Further research is needed to establish the robustness of the procedures in realistic settings, specifically for the differential measurement error and the alloyed gold-standard problems, and for determining optimal designs for selecting a validation sample.

We believe that these statistical methods for "bias correction" are ready to be used in case-control studies in some limited situations. In particular, RCE can be an efficient tool when measurement error is small and the error structure is reasonably well understood. If the measurement error is large or the error structure is not known—as often is the case in practice—a semiparametric estimator can be used as a robust alternative, at least when there is no important differential measurement error. For additional economy, an efficient design using stratified sampling should be considered as a way to select the most informative subjects in the validation sample. Evaluation of the performance and utility of these procedures will require several applications in studies with hard-nosed critiques of the validity of the underlying assumptions.

Validation studies, whether or not designed for bias correction, can be crucial to increasing the value of epidemiologic studies. An instructional example is the rapid progress in understanding the epidemiology of cervical neoplasia that followed the identification of polymerase chain reaction as a sensitive and specific assay for infection with oncogenic human papillomavirus[14] through intra- and interlaboratory studies of replicability.[15] The best way to reduce bias from measurement error is to improve tools for measuring exposures including biological markers, environmental samples, and questionnaires.

References

1. Sturmer T, Thurigen D, Spiegelman D, Blettner M, Brenner H. The performance of measurement error correction methods for the analysis of case-control studies with validation data; a simulation study. *Epidemiology* 2002;13:507–516.
2. Kipnis V, Freedman LS, Brown CC, Hartman AM, Schatzkin A, Wacholder S. Effect of measurement error on energy-adjustment models in nutritional epidemiology. *Am J Epidemiol* 1997;146:842–855.
3. Wacholder S. When measurement errors correlate with truth: surprising effects of nondifferential misclassification. *Epidemiology* 1995;6:157–161.
4. Wacholder S, Armstrong B, Hartge P. Validation studies using an alloyed gold standard. *Am J Epidemiol* 1993;137:1251–1258.
5. Neale RE, Green AC. Measuring behavioral interventions by questionnaires and prospective diaries: an example of sunscreen use. *Epidemiology* 2002;13:224–227.
6. White JE. A two-stage design for the study of the relationship between a rare exposure and a rare disease. *Am J Epidemiol* 1982;115:119–128.
7. Breslow NE, Cain C. Logistic regression for two-stage case-control data. *Biometrika* 1988;75:11–20.
8. Holcroft CA, Spiegelman D. Design of validation studies for estimating the odds ratio of exposure-disease relationships when exposure is misclassified. *Biometrics* 1999;55:1193–1201.

9. Weinberg CR, Sandler DP. Randomized recruitment in case-control studies. *Am J Epidemiol* 1991;134:421–432.
10. Wacholder S, Weinberg CR. Flexible maximum likelihood methods for assessing joint effects in case-control studies with complex sampling. *Biometrics* 1994;50:350–357.
11. Langholz B, Borgan O. Counter matching: a stratified nested case-control sampling method. *Biometrika* 1995;45:69–79.
12. Breslow NE, Chatterjee N. Design and analysis of two-phase studies with binary outcome applied to Wilms tumor prognosis. *Appl Stat* 2002;48:457–461.
13. *S-Plus* [computer program]. Seattle: MathSoft, Inc., 1996.
14. Schatzkin A, Freedman L, Schiffman M. An epidemiologic perspective on biomarkers. *J Intern Med* 1993;233:75–79.
15. Schiffman M, Herrero R, Hildesheim A, *et al*. HPV DNA testing in cervical cancer screening: results from women in a high-risk province of Costa Rica. *JAMA* 2000;283:87–93.